CHAPTER 5: COLLECTING THE DATA

Data is information that can be used to create statistics, graphs, and tables to summarize your evaluation findings. But data is more than simply numbers. Data can come from written questionnaires, interviews, record counts, measurements, group discussions, and even pictures and story telling. And data collectors will vary in their ability to successfully get persons with information to participate. So, careful selection of the evaluation team's data collection staff is important.

Data collection can be accomplished through a variety of planned actions designed to gain the insights needed to answer the evaluation questions. Working from the program consensus maps and logic model, you will develop your set of prioritized evaluation questions to then guide the planning of the data collection process and methods. The process will be completed when you have determined mechanisms for collecting the essential data/information needed.

Once collected, data/information will need to be processed, computerized, edited to eliminate errors and assure data completeness and accuracy (to the degree possible), and then analyzed. Once analyzed, the results still have to be interpreted and incorporated into an evaluation report that can then be disseminated to the stakeholders.

The Steps For Data Collection

The steps for data collection include:
- Clarifying what data/information are needed;
- Reviewing data collection resources;
- Locating data sources;
- Choosing the most appropriate data collection methods and instruments;
- Deciding how much data is enough;
- Deciding who will collect the data;
- Deciding what sources the data will be collected from;
- Obtaining "informed consent" from clients willing to allow their data to be used in the evaluation process;
- Deciding when and where the data is to be collected;
- Protecting data confidentiality during data collection and processing;
- Determining how, where, and by whom data will be processed and analyzed; and,
- Staying attentive and sensitive to what the evaluation process is revealing.

A participatory evaluation process will assure the use of culturally appropriate ways to collect data from your community. If an outside evaluator is involved, be sure that the data collection methods proposed by the evaluator are reviewed by and agreed to by your community partners. Remember, the quality of your evaluation findings rests on the quality of the data you collect. An inclusive participatory process assures highest quality without sacrificing cultural appropriateness and practical usefulness of the findings.

Common Data Collection Challenges

The following questions illustrate common challenges involved in data collection:

- *How Complex Should Our Data Collection Process Be?* Data collection can be simple or complex, depending on what you are interested in finding out and the available resources. Unless you are required to meet scientific research standards, keep your evaluation methods as simple as possible.
- *Where Do We Get Data?* Often, the most effective data collection techniques are the ones you already use. Identify those sources of information that are available and can be used to evaluate program operations, client outcomes, or community impacts. As you begin collecting information, remain flexible enough to modify your process to assure that the findings are optimally useful for enhancing program operations and outcomes. If you discover critical issues as your evaluation proceeds, you will want to note these and enhance your data collection efforts to provide sufficient insight into each issue for sharing with the program management team. At the same time, you do not want to become sidetracked.
- *How Much Data is Enough?* Collect only that information/data that is needed to answer the evaluation questions agreed to by the stakeholders. Do not gather data that you won't use. And, prioritize the questions to be answered and focus your efforts on these.

Clarifying What Types Of Information You Need

Think about what kinds of data will best answer your questions -- "quantitative" or "qualitative" data. The strongest evaluations collect both.

Quantitative methods produce things that can be counted or measured and are most useful when we want to know: How many? How much? Rank order? Degree of change? Projected program impacts? The varieties of quantitative data types are important because they limit or enable the analyses that can be performed. The three primary types of quantitative data are:

- *Nominal Data* – counts by categories that cannot be ranked.
  For example: "Yes" versus "No"; "Male" versus "Female"; or, "Hispanic" versus "Non-Hispanic".
- *Ordinal Data* – counts by categories that can be ordered or ranked.
  For example: "None", "Moderately", "Most"; "Don't Smoke", "Occasionally Smoke", "Smoke Daily"; or, school grade.
  (Ordinal data is useful for assessing relative changes over time.)
- *Interval Data* – measured or scored data that falls along a scale that produces measurements that allow the observer to calculate the measurement intervals.
  (For example: temperature; weight; age; blood pressure; or, test scores.)

Of the three types of quantitative data, interval data is the most powerful because it allows comparisons of measurements over time so the degree of change can be credibly assessed. Interval data enables the most powerful statements to be made regarding program outcomes and impacts. The downside of interval data is the cost that can take to collect it. If resources are an issue, you might reserve interval data measurement for answering only the highest priority evaluation questions.

Qualitative data, on the other hand, often includes quotes from interviews or from the descriptive stories people share that bring the culture, spirit and vibrancy of your program alive. The purpose of qualitative data is to:

- Capture contextual information.
- Provide depth and detail for understanding and interpreting observations.
- Discover the meaning and impact of your program from the perspectives of others.

"A Story From The Hispanic Culture"

My mother worked in a hospital in a town where I grew up, a little town called Miami, Arizona, a copper mining town that was about one-half Mexican and one-half White. There were some Indians on a reservation twenty-four miles from there, but mostly Hispanic and White, and most of us poor. She worked in the kitchen and the nurses there, most of them were non-Spanish speaking, but we used to get a lot of Spanish-speaking people in the hospital. So, they went to my mother and asked her if she would help them with some basic phrases they could read into the cases. So, she had written them down and even done them phonetically so that they could get people to cooperate with the blood pressure, temperature, things like that.

Well, the nurse came down and she said, "I need some help upstairs," and my mother went to do the translations. My mother said to the nurse, "This lady understands what I'm saying. What were you telling her?" And the nurse said, "Well, I read the stuff just like you told me." So my mother turns to the patient who doesn't speak English and said, "This lady was talking to you in Spanish." And the lady said, "Oh, she was? Well, I was listening in English."

Now, there is an important message in that. It's like, she was trying to use mainstream frame of reference because that's where she was. And mainstream frame of reference, for a change, was trying to reach her, so there was definitely a misfire there."

Told by C.H. Balderama
CENTERED Blue Ribbon Panel Member

Data Sources

Now that you know what you would like to learn and your resource situation, you need to determine who has the data you need. There are two kinds of data sources: "primary data" that you collect yourself, and "secondary data" that someone else has collected but allows you to use.

Secondary Data:

What data will be most helpful to you depends on the data/information necessary to answer your evaluation questions, and which data sources are most credible. While secondary data is usually less expensive and quicker to obtain, it must meet your needs. Consider the following questions to help you judge how "good" the data source is:

- Whose data is it? How was it collected? Who funded it and for what purposes?

- How accurate is the data? How up-to-date? Do you have a mechanism for validating the data?

- How well does the data reflect the community? Is it representative?

The answers to these questions will enable you to identify possible sources of bias that you need to be aware of and reflected in your evaluation report if you choose to use them.

Chapter 8 includes a listing of secondary data sources that may be helpful to you. You can also request data from your state and/or local health department. You will want enough data to establish baselines for the program indicators you've chosen so you will have a point of reference to compare with over time as your program progresses.

Remember, not all data sources are going to provide data useful for your purposes. You will need to determine:

- Does the secondary data source represent the target area for our program?
- Does the secondary data source allow data to be analyzed in accordance with your program's priority populations (i.e. racial, gender, and age categories)?

If the data source does not allow you to approximate the boundaries of your program's target area, it will probably be much larger than the area that your program has defined as its target. This means that whatever impacts your program expects to make in enhancing the health of the community will be hidden because the secondary statistics include data from areas in which your program has not been providing services. If you still decide to use this data source for assessing program impacts, it will almost always take longer to evidence those impacts because the data includes persons not involved with the program. If you can, select target areas with boundaries that coincide with secondary data sources, such as using census tract or county boundary lines.

A common problem with secondary data is that the source does not summarize its data into the same categories as you need for your program. For assessing your program's impacts on local health disparities, you will want data that is categorized specifically into those racial/ethnic subpopulations that your program targets. Data broken down only into general "White", "Black", or "Other" categories is not helpful. For example, if you are working with American Indian, Hispanic, or Asian/Pacific Islander communities that have been lumped together into the "Other" category, how could you possibly develop baseline prevalence rates or trends in incidence rates over time?

Because of the *Healthy People 2010* national initiative, efforts have been initiated to make data available for those subpopulations that are the focus of the national initiative.

However, this will take time as data has not traditionally been collected and reported to this level of detail in many juridictions. Just placing your request to the secondary data sources for data broken down to the specific racial/ethnic group level that you need to best serve these high interest populations will strengthen appreciation for the need and the likelihood to collect this data. So, strongly express your need for your specific data needs and encourage others to do the same.

Primary Data

The alternative to using secondary data is to develop your own data — primary data. In this case, you control what is collected, from whom, and what is done with it. This can be highly valuable data, because it is specific to your target population. However, primary data also has its own challenges and limitations. Consider the following:
- What information already exists about how your program and how it is working?
- What stories have you heard from clients, staff, and others about the program?
- How does the program track how many clients receive program services?
- How does the program assess client satisfaction with the services offered?
- How are you assessing changes in client knowledge, attitudes, or behaviors as a result of program participation?
- What information did you use to support the need for funding for the program?
- What changes in client health outcomes or in indicators of community health have you targeted as program objectives/goals and how are you monitoring progress?

Also, think about who may have information that would be useful to your evaluation process:
- Program clients (past, present, and drop-outs);
- People who chose not to participate in your program at all;
- Program staff;
- Community members who know the program;
- Family members or friends of program clients;
- Others working who have witnessed client changes attributable to the program;
- Staff working in other agencies or institutions you are trying to make changes in;
- Community members able to observe changes at the community level; and/or,
- Key authorities (e.g. Board of Health, School Board).

It is extremely important to include community members most impacted by racial and ethnic disparities. Their voices are critical, yet they are often not included. As you collect information from the community you are developing your own primary data sources. You will have the opportunity to gain insights that your evaluation would miss if you only focused on those directly involved with the program. Remember, the community can help identify the reasons your program may be underutilized by certain target groups; and, help your staff to be more sensitive and responsive to those groups.

Many community programs operate under challenging conditions where racial inequities and other negative social forces work against efforts at social progress. If this is the case in your program's target community, you need to collect data from those within the community willing to share their perspectives. If and when the community raises these

issues, they should become high priorities for the program to add to the evaluation. Beyond assessing the perspectives of the community, if action on the issues lies beyond the scope of your program, then you should still refer the issues to appropriate parties/agencies with cc's to the community so they are informed of those referrals. You may still want to incorporate these concerns into the periodic evaluations to demonstrate your support of community interests.

Choosing Data Collection Methods And Tools

The consensus mapping process (Chapter 2) guides you in developing the questions to be answered by your evaluation. Some can be directly answered by a count or a measurement, while others may require interviews or surveys. Still others may best be answered by culturally appropriate group discussions, such as took place during the consensus mapping process.

Data can be collected in a variety of ways, including:
- *Direct assessment:* Counts or measurements.
- *Survey:* Standardized or open-ended questions (for contextual insights).
- *Interview:* Structured (guided by predetermined questions); or, unstructured (open-ended responses).
- *Focus group:* Structured and led by a neutral facilitator to address a predetermined set of issues.

The ability to collect the needed data depends not only upon selecting the appropriate format, but in the case of interviews, surveys, and focus group discussions, it also depends on the ability to raise the issues of interest in a manner that encourages honest and full responses. This process involves working with informants on the wording to be used, and how questions should be framed to be most acceptable.

Respondents can also guide the selection and training of data collection staff, to optimize their reception by the community. Whenever possible, persons used for data collection should come from the community. All data collection instruments should be pilot tested with members of the target community prior to being finalized. Sometimes the wording and the order in which questions are asked can make a difference, and the pilot testing will help to identify any need for refinements.

Resources that can provide more in-depth discussions and examples of data collection instruments are listed Chapter 8. Invite an experienced evaluator from the health department, university, or another CBO to help as you develop your data collection instruments. Table 1 reviews common data collection methods. You may decide to use less mainstream methods, but the funder and/or independent evaluator should be involved early to assure the acceptability and reliability of the method—does it measure what it is supposed to measure? Consider the following options to overcome concerns about reliability:
1. Conduct parallel data collection systems and compare the results.
2. Reinforce use of culturally appropriate methods with validity documentation.

3. Invite concerned community partners to observe data collection efforts so they can see how people are responding to the method(s) used.

Once you have settled on which methods to use, you need to develop your data collection instruments. This includes development of both the interview guide (for training data collection staff), and data collection instruments. You can develop these or borrow ones that others have found to be useful.

Table 1 . Commonly Used Data Collection Methods.

METHOD	ADVANTAGES	DISADVANTAGES
Written Survey	Comfortable for people coming from a reading and writing culture. Can collect data from many people at the same time. Relatively inexpensive. Can be anonymous if sensitivity is an issue.	Unless supervised, people may not return the survey. Literacy issues can interfere with response rates. Follow-up for fuller explanations of responses may not be possible. Incomplete information is common. Little chance of building relationships with respondents.
Interview (individual or focus group)	More comfortable for people from an oral culture. Interviewer can clarify issues for participants. Reduces literacy issues. Can develop evaluator-participant relationship.	Time-consuming, labor-intensive, relatively expensive. Evaluator training is critically important. Data may be hard or very time-consuming to analyze (qualitative).
Records Review (data collected for other purposes)	Unobtrusive. May enable validation of data collected through other methods.	Time-consuming, expensive, & labor intensive. Relies on accuracy of the record and on knowledge of data collector. Incomplete information is common.
Direct Observation	Opportunities for new issues to arise. Provides context. Observer sees first-hand what the community or program is like and the intensity of feelings expressed by both clients and program staff as they interact.	Time-consuming, expensive, & labor intensive. Observer training is critical as the presence of the observer may cause clients or staff to change their normal behavior. Can be open-ended if the process lacks clarity regarding what the observer is to look for and how observations are to be interpreted. Data may be hard to analyze.
Clinical Test	Stakeholders value highly. Objective assessment is possible; enables detection of changes in the selected	May violate client confidentiality and require permission to view results. Evaluator training or external expertise may be necessary

		indicators being measured. Results can be validated.	to interpret results. Time, labor, & technically intensive; expensive.
Knowledge, Attitudes, & Behaviors Surveys (a.k.a. KAB Surveys)		Can be standardized. Stakeholders may value highly. Provides baselines to assess change over time.	If not culturally appropriate, may provide misleading results. Literacy may impact results. Self-reports are difficult to validate.
Storytelling		Common in some cultures. Allows participants to guide the conversation.	Open-ended. Difficult to analyze. Produces data on a variety of topics, some unrelated to the topic at hand.
Writings (journals, logs, other writings, etc.)		Good for reflection-oriented culture or communities. Captures what someone "says". Informer convenient. New or embarrassing ideas able to emerge "safely".	People must write comfortably. Hard to analyze as very open-ended. Difficult to obtain as it depends on writers "sharing" writings. Knowledge of "shared" writings may influence what is written—self conscious censoring.
Drawings, Photographs		For visual cultures & communities. Avoids literacy issues. Provides a non-word, non-number method for sharing ideas and feelings.	Difficult to analyze, may need explanations in writing or verbally. May be difficult for some to understand or find credible. Very open-ended.

What Questions to Ask And How To Ask Them

Your stakeholders represent the community's interests and help to inform it about the evaluation and why it is in its interest to support the evaluation, and encourage community cooperation. The questions to be asked must be sensitive and appropriate to the unique culture and history of the community. The wording of the questions must show respect and appreciation for those willing to share the information sought.

> **"AN ALASKAN NATIVE STORY: THINGS PEOPLE DON'T TALK ABOUT"**
>
> While working in Anchorage on an evaluation project for a cancer screening program for Alaskan Native women, we wanted to better understand the following: What are the cultural beliefs that these women have around cancer? We were told, "You can go in and ask even the nurses and some of the community health aides, and people who deal with them about cultural beliefs. But, they're not going to want to tell you because they have the beliefs themselves and when you talk about them, you give them power. When you say these things out loud, you're going to give the cancer more power," and there are things that people don't talk about. So as an outside evaluator coming into a community, you can't even say, "Oh, we have to be culturally sensitive and collect information on some of the stuff." You're not even going to get it, you're not going to get people talking about it.

Ask questions on subjects that are appropriate for the community, but beware of introduction of biases by the way questions are worded. For example, do not ask questions in a way that places guilt or reflects negatively on the informant. Let's say that, based on participants' age, weight, and overall health situation that doctors with your program are recommending selected participants walk regularly to increase their exercise. Now consider how the participant might feel if their community has no sidewalks – let alone walking trails – to enable getting out daily to walk on? The following are examples of evaluation questions that raise the exercise issue in different ways:

- What did your doctor tell you about why you should get more exercise?
- To what degree do you agree with your doctor about the need to get more exercise?
- What was helpful in terms of your doctor's encouragement?
- What would have been more helpful?
- What would support you in getting more exercise?
- How might you need to change?
- How might other things need to change?

The close-ended first question can produce "guilt" in the participant without yielding information that could help in finding ways to make routine exercise more likely. The open-ended questions that follow allow issues to be raised that could enhance the success of the client and your program, if you could convert the insights into recommendations to be shared with other community partners.

If you use methods that ask open-ended questions, it is helpful to do this in ways where a neutral facilitator can guide the conversations or interviews to keep the discussions focused on program-related issues, but without killing opportunities for important unanticipated issues to surface.

Sampling: How Much Data Is Enough?

In general, it is not possible or necessary to collect data from every member of the community or even from all clients in order to evaluate the program. When done properly, sampling can provide sufficiently accurate insights to satisfy your needs while keeping the effort and cost of data collection to a minimum. The minimum sample size you will need to use will be determined by your study design, the nature of the questions you are seeking to answer, and the precision of the results required by your stakeholders. An epidemiologist or statistician can help you calculate the minimum sample sizes needed so you can be sure that your sample sizes are adequate for your intended purposes, while controlling costs.

The following are important rules to consider when sampling:
1. The population to be sampled must be well defined and the members of that population reliably identifiable.

2. The sampling mechanism must be such that for every member of the population the probability of that person's inclusion in the sample is known and is not zero.
3. The sampling design (which is based upon these probabilities) must be implemented as planned (or bias will be introduced).
4. Conclusions based on the sample results can only be attributed to the sampled population.

Sampling Options

- *Random Sampling:* Produces results representative of the population from which the sample was taken. The probability of being included in the sample is the same for all individuals in the population. The goals and objectives of the study dictate those to be included in the sample.

- *Stratified Random Sampling:* A sampling fraction is determined for each strata based upon the relative proportion of that strata within the total population to be sampled. If this population is made up of two subpopulations (African Americans, 30%; and, whites, 70%), those selected would be included in the sample in proportion with their representation in the population (i.e. if 1000 persons were to be included in the sample, the goal would be to include 300 African Americans and 700 whites). This is called a "proportionate allocation" sample.

  If the sample contains equal numbers of each subpopulation (i.e. 500 African Americans and 500 whites), regardless of their proportion within the total population, then this is called an "equal allocation" sample.

  If the random sample is a true representation of the population as a whole, it is called a "self-weighted" sample.

- *Cluster Sampling:* Groups rather than individuals are selected for inclusion in "cluster" sampling. Random selection still occurs, but groups are selected at random and all individuals in the selected groups are sampled. When the data is analyzed, the cluster is the unit of analysis rather than the individual.

- *Systematic Sampling:* The first individual is selected at random, then all other selections are made based on a predetermined pattern. For example, every 10th person might be selected until the desired number needed for the sample is reached.

- *Purposeful Sampling:* In some cases a random sample will not meet the needs of the evaluation – for example, if you are interested only in members of the Catawba Nation (a federally recognized American Indian tribe), then you might first screen on the basis of race/ethnicity, then from these, sample only from those who report being Catawba. In these cases you purposefully select *those key individuals because* they have the information you need. When using this sampling method you need to understand that while you have gained specific insights about the key group sampled (Catawba), you cannot make generalizations from this group back to the larger population group (American Indians) as you could with random sampling. Usually,

purposeful sampling is done to gain additional more detailed insights into high interest subpopulations.

- *Convenience Sampling:* Sampling by selecting anyone willing to be surveyed or to show up for a group interview/focus group produces the least credible data and is frowned upon by most funders and public health professionals. One use of convenience sampling is to field test survey instruments prior to finalizing the documents for actual use.

No matter how you select the sample of persons from whom you will collect your data, you will need to be able to explain to others the sampling method(s) used. This allows others to appreciate the strengths and/or limitations the sampling method(s) may have imposed. Sampling methods do impact the quality of the evaluation results and must be taken into consideration as the data results are interpreted. As previously suggested, you may want to involve an epidemiologist or statistician to help in: considering your sampling options; calculate the minimum sample sizes needed for your intended purposes; and, fairly describe the methods used and limitations imposed by those methods.

Who Will Collect the Data?

One way to enhance the quality of the data you collect is to use locally recruited data collectors who understand the local "culture" and can encourage active and honest participation. Persons selected in the data collection sample need to feel that the insights they provide are valued, and that anything they share will not result negatively on them. Trust comes more quickly when the data collector is from the community.

You and your stakeholders should decide who would make the best data collectors, because you know both your community and your program. If there is a need to control community biases from impacting data collection, then it may be best to use data collectors who come from outside the community. You and your partners need to weigh the advantages and disadvantages of your options.

If you have a diverse data collection team, you should make the most of the strengths and experience of each team member in deciding roles and responsibilities. Training and supervision of the collection team is essential for achieving high quality data collection.
Think about offering payments or the use of alternative incentives for rewarding the community data collectors you employ. While skill-building or "giving-back-to-the-community" incentives might be enough, cash payments may be a more suitable way to acknowledge people for their sustained commitment to supporting the evaluation process. Whenever possible, community partners should be paid for their important contributions to the evaluation effort, and acknowledged publicly for their supporting efforts to eliminating the health disparities that exist within their community.

When Is Data To Be Collected?

You will need to collect data before your program begins in order to establish baselines for the indicators that you will be monitoring throughout your program. You will then re-assess these indicators periodically throughout the program, and then again at the end of the program. This will allow you to determine short-, medium-, and long-term client outcomes and the community-level impacts. You will also want to collect data throughout your program to monitor the operational processes and be able to make and assess program refinements.

Informed Consent: Protecting People Who Provide Data

You also need to consider what types of information those providing data will need in order to make an informed choice as to whether or not to participate in the evaluation process. While it is rare, information you collect may be subject to subpoena. Therefore you need to consider such options as collecting data anonymously and/or avoiding questions that can put those who provide data to you at risk. Confidentiality protection must be included in training for program staff and for all members of the data collection team. Confidentiality protection assessments should also be incorporated into the evaluation. The National Institutes of Health provide a number of excellent on-line tutorials about confidentiality and human rights that you may want to use in training your staff. Table 2 summarizes some of the data confidentiality issues that you will want to address as you plan your program evaluation.

Table 2. Data Confidentiality And Data/Client Protection Issues.

QUALITATIVE DATA	*QUANTITATIVE DATA*
Includes notes taken by hand, tape recordings, typed transcripts, and reports.	Includes data collection instruments, databases, compilations of findings, and final reports.
-Who is allowed access? -How can you control access? -How will you share stories while protecting the identity of the source? -Consider substituting names and other identifying information with codes. -Who owns the data? -How can it be used after the report is completed? This should be made clear to outside evaluators, staff, stake-holders, and to those contributing their stories, interviews, or who are under observation.	-Who will have access to the raw data? -Who will need client identifiers to enable validation? -Where and how will data be kept and secured? How will data be shared without revealing the identity of the source? -Consider using codes or unique identifiers instead of names. -Who owns the data? -How can it be used after the report has been completed? This should be made clear to outside evaluators, staff, stakeholders, and to those contributing their data.

Always secure any records that may contain client or patient identifiers. Ideally records should be secured inside a locked file cabinet inside a lockable room. Only those staff

with a valid need-to-know should have access to those records. A responsible person should be charged with maintaining the records security and for supervising any release of records. Conditions of release should be addressed in the program's Institutional Review Board (IRB)-approved data release policies.

Here are some examples of things you will need to bear in mind about each person providing data (informant):

- Who they are? Do I need their name and other personal identifiers, or can I substitute a unique code and still satisfy the program's needs?
- What is their role in the program and/or the evaluation? Client? Staff?
- Why is this data/information being requested?
- Why is it important for this person to participate?
- Who else will have access to the data?
- How will the data they provide be used?
- What are the consequences if they chose not to allow use of their data?
- What would happen if they decide after agreeing to allow their data to be used, to change their mind? What impacts will this have on them? On the program?
- What types of data will be collected? On what topics?
- What kinds of things would have to be reported to authorities (e.g., child abuse) if they came out during data collection?
- What are the risks for those clients/staff members who agree to participate?
- What are their benefits?
- What incentives are you providing?
- How much time will this take?
- What will happen to the data/information once the evaluation is completed?

Make sure participants understand that *they may ask questions at any time* of a responsible person familiar with the program. You should *document* in writing or on videotape that *an official informed consent has been obtained* from each person involved in providing data/information collected during the evaluation process.

Incentives And Stipends

To encourage clients or community members to participate in the evaluation process you may want to offer some sort of incentive. Cash stipends, a gift coupon, clothing, or other incentive may be offered in appreciation of their participation. The nature and scale of the incentive is best decided with both the community stakeholders and funder(s). If there are expenses that will be incurred to support the evaluation process, these need to be anticipated before they are incurred, or reimbursement procedures established and described ahead of time. Non-monetary incentives may include public recognition. In some cases, giving persons or organizations something they need may be the preferred incentive – e.g., providing training or other technical support that your resources allow. In other cases, just knowing they are participating for a worthy cause or being listened to may be sufficient.

Managing Your Data

A data flow plan should be developed that illustrates: Where data is collected; how it will flow to computer entry; and, where data will be analyzed. The data flow plan should show where and how data coding (if used) will take place, by whom, and where the code key will be located and secured. The plan should include a detailed description of who is responsible for data security, including how access to the code keys (if used) will be controlled and the criteria for breaking a code (subject to Internal Review Board (IRB) approval).

Data storage, editing, transformation, reproduction, dissemination, and disposal also need to be considered relative to data security and ownership issues. Data quality assurance systems need to be developed, including data management protocols and how monitoring for protocol adherence will take place. The data flow plan and protocols need to address when and how logical data edit routines will be used, and how missing data/information will be handled. Planned data analyses should be described, including designation of the person(s) responsible for conducting those analyses and interpreting the results.

Finally, the ultimate fate of the data collection documents and data must be addressed. Will paper records be retained or destroyed? If retained, will the records need to be de-identified (removal of names and other identifying information), or will coded data sets be retained and need to be secured against future abuse?

By setting up and describing the data flow systems and ways of operating in advance, you help assure that your data is consistently recorded and organized in ways that insure both data quality, and data confidentiality. Thorough training of data collectors and data handlers; judicious use of back-up systems (e.g. two tape recorders for interview sessions in case one breaks down); backing-up of all data bases with secure off-site storage (in the case of a fire, break-in, etc.); and, continuous quality assurance monitoring for staff adherence to program policies and protocols are all essential elements of the data plan.

Data Management Supervision

Supervision of all aspects of data collection and processing is a major management responsibility since the data will eventually be used to document the activities, outcomes, and impacts of the program, as well as enable cost benefit assessments to be made. The date analyses produce the results that ultimately determine the fate of the program and of the services upon which the clients and community may have become dependent. This is a major responsibility that also requires periodic evaluation to assure its effectiveness.

You are encouraged to hold regular meetings with those involved in data flow so they can share their experiences and frustrations and help in the early identification and resolution of potential data-flow problems. Be prepared to refine your data collection instruments as you learn what works and what doesn't. Always seek to do what appears to be most effective for accomplishing the stated objectives and goals. When refinements involve elements needed for meeting a stakeholder's needs, be sure the stakeholder is involved. Keep in mind that it is important that your data be organized to enable others to analyze the same data using the same techniques in order to see if they produce the same findings.

Although interpretations may vary, data analysis should strive to achieve and maintain scientific rigor, including an objective description of any limitations due to data collection methods that might influence interpretations of analytical results.